

# 基于深度学习的光学遥感机场与飞行器目标识别技术

牛新, 窦勇, 张鹏, 曹玉社

国防科学技术大学并行与分布处理重点实验室, 湖南 长沙 410073

## 摘要

机场与飞行器目标识别是遥感数据分析中的典型应用。研究了光学遥感大数据环境下面向机场与飞行器目标识别的深度学习技术。为此, 构建了一个面向高分光学遥感图像的机场与飞行器目标秒级识别系统。使用迁移学习的方法在有标签样本稀缺的情况下有效构建深度网络, 利用目标先验知识对潜在目标进行高效提取, 并提出一种层次式的级联深度网络识别架构, 实现“大范围、小目标”的实时识别。实验结果表明, 采用相应技术, 基于深度学习方法可以在秒级时间得到比传统方法更高的识别精度。

## 关键词

光学遥感; 目标识别; 深度学习

中图分类号: TP183

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2016054

## *Airport and flight recognition on optical remote sensing data by deep learning*

NIU Xin, DOU Yong, ZHANG Peng, CAO Yushe

National Key Laboratory of Parallel and Distributed Processing (PDL), National University of Defense Technology, Changsha 410073, China

## *Abstract*

Airport and flight recognition are the typical remote sensing applications. For the big optical remote sensing data, deep learning techniques for airport and flight recognition have been studied. To this end, a seconds' response airport and flight recognition system for optical remote sensing data was built. To obtain effective deep learning model with limited labeled samples, transfer learning approach has been employed. Prior knowledge has also been explored for efficient object proposal. To achieve real-time performance for such recognition with "large region and small targets", a cascade framework of deep networks has been proposed. The results of experiments show that, by the proposed deep learning approaches, significant improvement on recognition accuracy could be achieved with seconds' response.

## *Key words*

optical remote sensing, object recognition, deep learning

## 1 引言

遥感技术是依靠飞行器、人造卫星等载体,通过光学、雷达等传感器对远距离目标进行探测和识别的综合技术。遥感技术已广泛应用于地图测绘、资源监管、环境保护、城市规划、地质研究、救灾抢险、军事侦察等领域。近年来,卫星、无人机等遥感技术得到了快速发展,遥感数据呈现出体量大、来源和类型繁多、更新和变化速度快等大数据的特点。根据美国卫星工业协会(SIA)《2015年全球卫星产业状况报告》<sup>①</sup>数据显示,截至2014年12月,全球在轨商业遥感卫星和军事监视卫星超过270颗。目前,航空航天遥感正向高空间分辨率、高光谱分辨率、高时间分辨率、多极化、多角度的方向迅猛发展。2012年我国发射的资源3号卫星每天的数据观测量可达10 TB,至2015年,其原始数据总量已达1 003 TB。而未来10年中,空、天、地平台部署的传感器每天获取的数据将超过10 PB<sup>[1]</sup>。遥感观测应用已进入大数据时代。

在遥感应用中,应用遥感图像对机场及飞行器进行目标识别一直是研究的热点。该技术已广泛应用于城市规划、航空交通以及军事侦察领域。在遥感图像上进行机场及飞行器识别,属于一种“大范围、小目标”的识别应用。受限于识别目标特别是飞行器目标的尺度,以往的研究主要集中于分辨率达到米级的高分辨率遥感图像,因此主要以航拍光学和高分辨率光学卫星遥感图像为主。在高分遥感大数据条件下,该应用突出困难在于以下两点。

- 目标的形态变化更加复杂:不同种类、不同排列角度、不同尺度、不同采集条件构成的目标不变性规律更加难以总结。
- 背景干扰更加明显:高分辨率条件

下,会出现更多与目标在某些图像特征上难以区分的背景目标,例如对机场目标而言,周围复杂的人造建筑区域或复杂路网都会造成干扰;干扰因素也更加突出,如对飞行器目标而言,停机坪上的纹理以及阴影变化都是很强的干扰。

为了充分利用大数据信息资源,进一步提高机场与飞行器目标识别精度和效率,需要高效的遥感大数据智能分析处理技术。近年来,一种称作深度学习的机器学习方法在大数据环境下异军突起,特别在诸多图像分析领域取得了突破性的进展,成为目前学术和工业界研究的热点。如何利用深度学习技术提高机场与飞行器目标识别的效果,也成为遥感应用领域关注的焦点。

## 2 光学遥感机场与飞行器目标识别发展概况

### 2.1 传统机场与飞行器目标识别面临的挑战

和其他目标识别方法类似,传统机场与飞行器目标识别方法一般也分为3个步骤,分别如下。

- 区域选择:在给定的图像上选择一些区域作为进一步识别的候选区域。
- 特征提取:对候选区域特征进行提取。
- 目标分类:通过已训练的分类器,依据提取特征对区域进行分类识别。

其中,目标特征的表达与学习是构建高效目标识别的关键。高效的目标特征应有较好的类间区分性与类内不变性,既能够辨识同一类目标中不同样本的复杂变化,也能对源于不同类的目标加以区分。

在候选区选取方法上,通常采用滑窗

<sup>①</sup> <http://www.sia.org/wp-content/uploads/2015/06/Mktg15-SSIR-2015-FINAL-Compressed.pdf>, 2015.

办法<sup>[2-6]</sup>,该方法将整个图像用不同尺度、不同长宽比的滑窗进行指定步长的扫描,所有滑窗扫描覆盖的区域作为候选区域进行进一步识别。这种方法可以极大地避免识别区域的遗漏,但计算量较大,有很多明显不含目标的区域也参与了计算,引起了不必要的开销,且步长、长宽比、尺度等参数也难以选择。另一种候选区域方法是基于图像分割<sup>[7-9]</sup>,即将整个图像根据内容的局部同质性先分割为图像块,这样目标就包含在某个或某几个相邻的分割块中,以此分割作为后续识别的基础。此种方法相比滑窗减少了很多识别候选,但通常候选区域不规则,需要制定的识别方法相对复杂。近年来,基于显著性的候选方法<sup>[7-9]</sup>得到很大发展,与滑窗和分割方法不同,基于显著性的候选方法通过图像的边缘、纹理等特征,找到目标区别于背景的所谓显著性特点,只选择图像中此显著性特征明显的地方作为候选区域,而忽略大多数背景。这种方法只提取少量有限的区域进行识别,可以极大提高识别效率。但是显著性特征的设计至关重要,在复杂场景下可能会因显著性区分度不高造成目标的漏检。

在识别方法上,关键是寻找合适的目标图像特征并设计相应的分类方法。传统方法多将特定目标识别作为特殊的通用目标识别应用,采用一些先进的通用特征,如HOG(histogram of oriented gradient,方向梯度直方图)、SIFT(scale-invariant feature transform,尺度不变特征变换)<sup>[2,4]</sup>等。但这些通用特征往往只能表达图像目标的底层特点,难以把握目标的不变性及与其他类型的区分性,在缺少训练样本的情况下,对于复杂场景难以取得稳定的效果。也有研究根据机场飞行器等具体目标特点,设计具有良好旋转、尺度不变性等特点的模板<sup>[10-12]</sup>进

行匹配,或根据目标的显著性特点等设计相应的专用手工特征<sup>[3,7]</sup>进行识别。但这些特征的设计依赖相关的专业知识,模型较固化,参数难调整,在面对新的复杂场景时,往往难以保持相对较高的准确性。

上述研究现状说明,在遥感目标识别中高效的特征表达与学习是识别的关键。除了挖掘目标的个体特征,其邻域特征或者说目标之间的相互关系特征的学习和应用也会对结果准确度的提高产生很大影响。而传统方法对于遥感数据特征的提取都存在一定的局限性。传统机器学习方法存在以下不足,一定程度上制约了其在大数据环境下的发展。

- 传统方法基于预设的特征模型准确性低,预设模型难以准确描述具体应用场景下的数据规律,且针对具体应用,模型的选择非常困难。
- 特征模型较固化,泛化能力差,难以掌握数据的复杂变化。大多模型本身固定,只可调整少量的模型参数。对于数据(如空间平移、旋转、尺度缩放等)的变化难以掌握。
- 由于模型较固化、自由度低,难以利用大量数据不断提高模型的表达能力。
- 大多数传统机器学习方法等价于浅层学习模型,难以提取目标的高层抽象信息。因此,高效的机器学习方法已经成为遥感目标识别研究中的热点。

## 2.2 深度学习目标识别及其在机场与飞行器目标识别中的挑战

深度学习是近年来在大数据环境下兴起的一门机器学习技术,其实质是通过构建具有多层结构的机器学习模型和海量的训练数据来学习数据中的复杂特征,用于之后的分析、预测等应用。由于其具备海量数据的学习能力和高度的特征抽象能力,

在诸多图像应用领域都取得了相比传统方法革命性的成功。

自2014年开始,基于深度学习(主要是卷积神经网络(convolutional neural network, CNN)<sup>[131]</sup>)的目标识别在自然图像应用中也取得了巨大突破。例如,以R-CNN(region with CNN features, 卷积神经网络特征区域)方法为代表的结合通用候选区域算法和CNN分类的目标识别框架(R-CNN<sup>[141]</sup>、SPP-NET(spatial pyramid pooling net, 空间金字塔池化网络)<sup>[151]</sup>、Fast R-CNN(fast region with CNN features, 快速卷积神经网络特征区域)方法<sup>[161]</sup>、Faster R-CNN(faster region with CNN features, 更快速卷积神经网络特征区域)方法<sup>[171]</sup>)。其候选区域采用Selective Search(选择搜索)<sup>[181]</sup>、EdgeBox(边缘框选)<sup>[191]</sup>或BING(binanzed normed gradient, 二值化赋范梯度)特征<sup>[201]</sup>等通用目标候选方法。而这些通用区域候选算法本质上还是遍历滑窗的方法。这些方法往往产生巨量的候选窗口(如在一幅遥感图像上往往可以产生数百万个候选目标区域),虽然类似SPP-NET、Fast R-CNN、Faster R-CNN等方法采用在卷积出的特征图上直接提取特征,避免了对每个候选区域进行大量的重复卷积操作,但由于没有减少非必要的目标搜索,其计算量依然很大。而且为了兼顾通用,目标的定位与框选并不十分准确,进而导致后期的CNN识别错误。虽然后期的算法包括Fast R-CNN和Faster R-CNN等都采用回归学习方法对目标的准确位置进行了修正,甚至有YOLO(you only look once, 框选识别统一)方法<sup>[211]</sup>、SSD(single shot multibox detector, 多层次框选识别统一)方法<sup>[221]</sup>等算法通过训练直接给出定位与识别结果,但这些算法都需要大量的学习样本用于预测准确的回

归位置,且大多设计为针对图像中占比大的目标,对于“大范围、小目标”的遥感识别效果并不理想。

受深度学习在通用图像领域的成功影响,机场与飞行器目标识别近年来也进行了一些基于深度学习的初步尝试。但深度模型的训练常常需要大量的有标签样本,受限于遥感目标样本采集的困难,大多数研究都集中于高分辨率光学遥感数据。这样可以利用光学图像的共性特征,在自然图像上通过大量样本预训练好的模型(通常是深度CNN),通过少量样本微调的方法引入光学遥感目标识别<sup>[9]</sup>,进而取得明显优于传统方法的识别效果。也有采用DBN(deep belief network, 深度信念网络)<sup>[231]</sup>无监督预训练的方法,直接在遥感图像上训练深度模型用于飞行器识别<sup>[5,6,8,241]</sup>。但使用DBN的方法,通常网络都相对深度CNN较浅,往往只有10层以下,模型泛化能力较弱,因此其高精度结果常常需要建立在高效的候选区上。同时目前的识别方法大多只能给出是否是特定目标的识别结果,而对目标类型进一步的精细分类缺少相应研究。

### 3 基于深度学习的光学遥感图像机场与飞行器目标识别系统与技术

针对遥感大数据的挑战,笔者构建了一个面向高分辨率光学遥感图像的深度学习目标识别系统。以机场识别和飞行器目标识别为例,提出了一系列方法期望解决复杂场景下“大范围、小目标”应用的快速准确定位问题,最终在Google Earth 19级图像上实现了高精度的机场、飞行器秒级定位。

该系统工作原理如图1所示,基于深度学习方法,首先进行机场定位,在机场范围

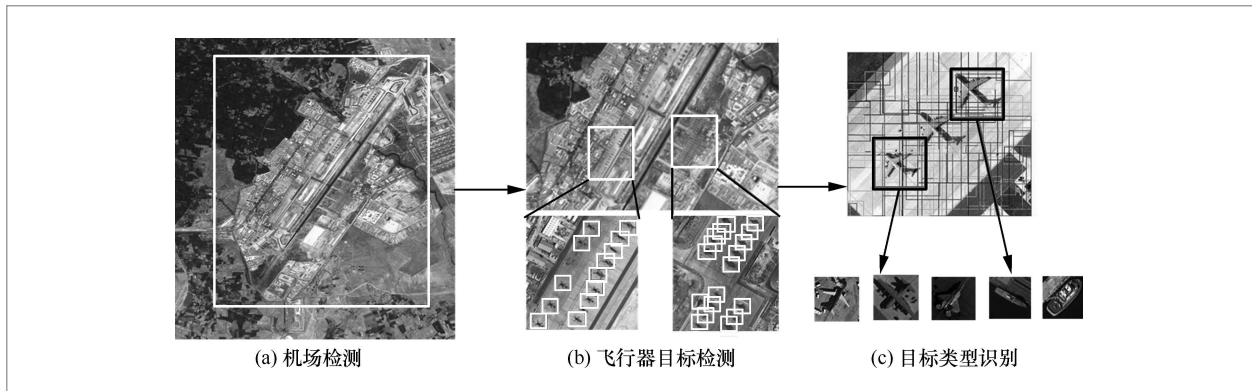


图1 机场、飞行器目标识别系统流程

内再进行飞行器的识别,之后再对飞行器的类型进行精细识别。在机场、飞行器目标识别系统中,依然采用候选窗口提取与分类器识别的识别框架:采用区域候选算法提取待识别目标区域;在候选区域上应用已训练好的深度学习模型,对区域所含目标进行分类识别。

如前所述,该类型目标识别任务的困难主要在于以下几点。

- 高精度深度模型构建与有标签样本相对稀缺的矛盾。遥感大数据大多数是没有类型标签的数据,特定目标样本的采集相对困难,而没有大量的标签样本难以对深度模型进行有效训练,例如机场目标,在中国范围内也仅能收集100多个样本。而对于飞行器样本,大量的人工标注依然难以收集足够的标签本来表现多种飞行器目标在不同环境下的复杂变化。可采集的样本数量与训练类似AlexNet<sup>[25]</sup>这样的深度网络所采用的具有百万量级的ImageNet图像库相比依然较少。

- 追求高分辨率的丰富信息与准确把握复杂目标特征的矛盾。高分辨率图像中,目标的变化更为丰富,背景干扰更加显著。虽然深度模型相比传统方法可以更有效地提取目标关键特征,但在样本采集困难的情况下,如何结合数据本身的先验知识,提升深度学习识别效果也是应用中需要解决

的问题。

- 大范围遥感图像中,采用复杂深度网络与追求快速识别之间的矛盾。深度模型往往较传统方法更加复杂,由于采用几十层到上百层的网络,计算量巨大。

因而,在大范围遥感图像中,对小目标采用深度学习模型进行识别,需要相应的高效系统设计。

### 3.1 面向光学遥感数据的深度卷积网络特征迁移学习

在图像分类与识别领域,CNN是目前应用最成功的深度模型之一。很多研究表明,较传统方法CNN可以学习到更高效的高层图像特征。但是训练一个识别率高、泛化能力强的CNN模型,往往需要大量有标签样本。目前多数深度学习模型是面向自然场景图像识别任务,研究比较充分,数据量庞大,而且都取得了非常优秀的成绩。而在光学遥感场景下,鲜有已经成熟的公开数据集。各研究组的研究分散,能使用的数据也是有限人力标注的少量数据,这对训练有效的深度模型构成了挑战。

尽管如此,CNN在两个任务有别、但数据结构相似的图像集上,依然可以采用迁移学习的方法,将大样本数据库上预训练好的模型应用在其他少样本应用中。在

机场和飞行器识别中,研究的光学遥感图像和自然场景图像在成像原理上相似,观测目标的颜色、纹理、形状等中、低层图像特征相近。因而,基于大样本自然图像预训练CNN,采用迁移学习方法,用少量样本加以进一步微调训练,可以有效构建面向光学遥感目标识别的深度CNN模型。

现有的CNN的网络结构,其底层卷积层大多提取图像的通用特征(边缘检测或者色块检测),中层和高层的卷积层或全连接层包含更多的任务相关特征。因此迁移网络时更关注的是网络顶层的迁移。直接有效的办法可采用网络微调的方式,将顶层的分类目标替换为现有任务目标,再用少量样本通过反向传播训练的方式对网络进行微调。训练样本的采集根据目标和场景的复杂程度而确定。一般采取细分类的办法,减少类内差别,提高识别精度,如将“背景”依据图像特点进行多种城区、农田、河流等细分类。同时,为了增强网络的顽健性,即增强网络对于目标在光照、平移、角度等复杂变化下的识别能力,对训练数据进行一定的变换和扰动扩充,在增加了训练数据量的同时可以

让网络充分学习到目标的变化,如在训练飞行器识别模型中,增加了飞行器训练样本7个方向的旋转变换、1种光照变换和1种镜像变换,使得样本扩充了 $(7+1)(1+1)(1+1)=32$ 倍,达到16万个样本的训练量级,如图2所示。

对于机场飞行器识别的深度CNN迁移学习,均采用在ImageNet上取得86%分类精度的AlexNet作为迁移源模型。在对6种机型和背景的分类问题上,采用迁移CNN的方法可以取得明显优于传统方法的效果,如图3所示(混淆矩阵中横坐标代表7种类型,其中1-6为6种飞行器,7为背景,色阶指示精度高低)。采用类似的方法,同样构造了10万样本量级的机场识别训练库。在对机场和背景共11类的分类结果上,也可以得到相比传统方法10%以上的精度提升。

### 3.2 结合先验知识的待检区域筛选

在机场、飞行器目标识别系统中,笔者依然采用候选窗口提取与分类器识别的识别框架。因此,除了目标特征学习,候选

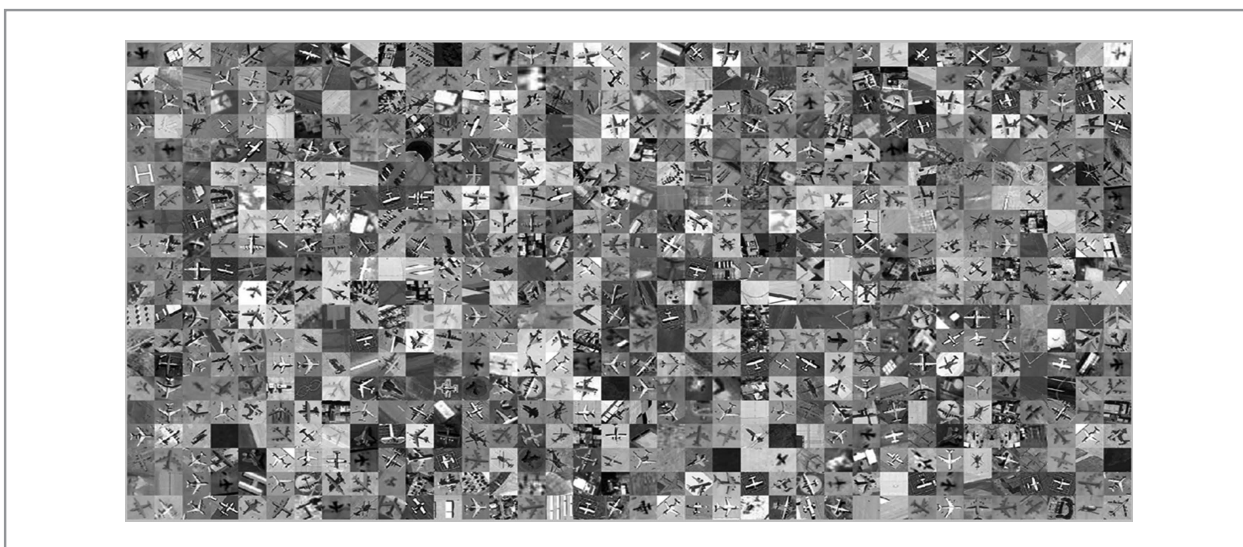


图2 光学遥感飞行器目标深度CNN迁移学习样本示例

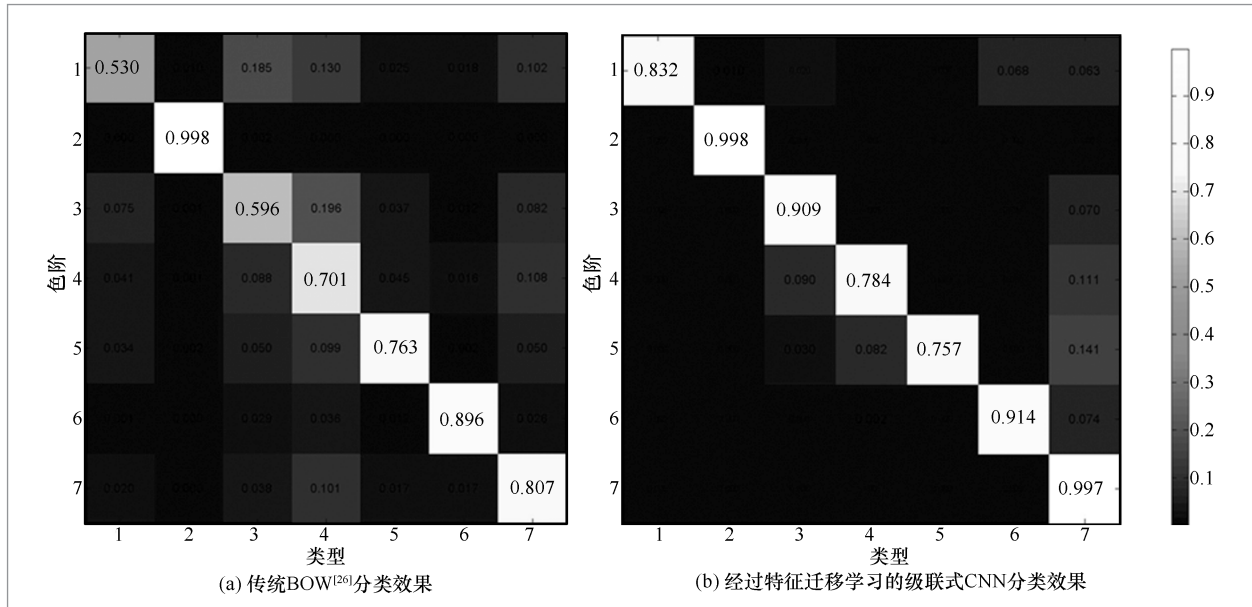


图3 飞行器细分类比较

窗口选取方法对系统识别效率有着极大影响。如前所述,常用的候选区域提取方法(如滑窗、BING、EdgeBox、Selective Search等)对于占图像比例较大目标的识别应用有良好效果。但在“大范围”的遥感图像识别中难以框选“小目标”的准确范围,且均会产生大量的候选窗口,对快速识别构成严重负担。而对于遥感图像中的特定目标,可以利用其先验信息,如机场的线性图像特征和飞行器的尺度及边缘分布规律,高效地构建候选区域。这些先验信息,还可以在后期识别中进一步提高基于深度学习的预测精度。

### 3.2.1 基于线段检测的机场目标候选区域提取

作为特殊人造建筑,机场跑道是机场区别于其他场景目标的主要特征。跑道一般可以被检测为具有一定分布规律的平行长直线段。据此先验知识,可以利用线段检测的方法提取有限且高质量的识别候选区域。具体过程如图4所示:使用LSD(line segment detector, 线段检测器)<sup>[27]</sup>算法快速提取初始线段;设计线段连接算法,恢复被截断的长直线段;依据线段之间的平行关系和垂直距离对线段进行分组,根据

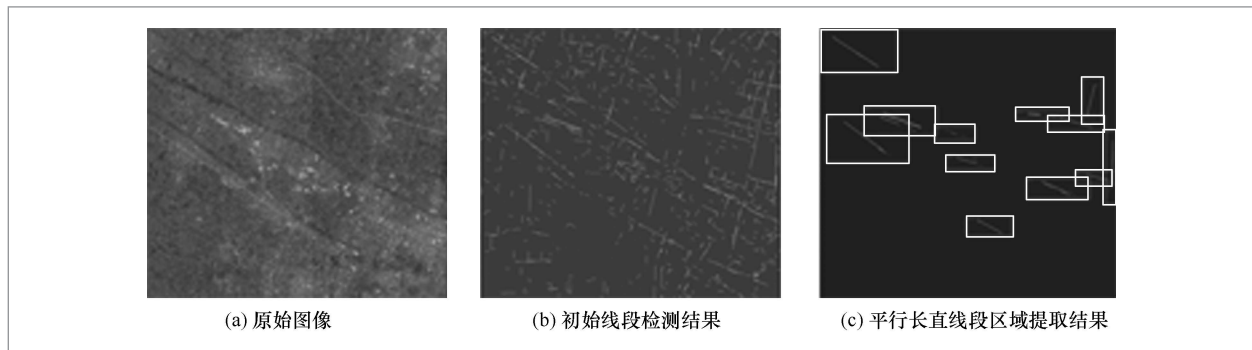


图4 基于边缘的机场候选区域提取示例

每一组线段的空间分布提取候选区域；为了进一步减少候选框数量，根据每组线段分布特征进行打分，筛除明显错误窗口。

经此方法得到的机场候选结果如图5所示（粗线框为机场真实区域，细线框为给出的识别候选区域），识别所需候选区域数量大幅降低，大多数案例中仅仅包含数个机场候选区域。同时，候选区域对目标的覆盖更加准确，从而可以有效提高机场识别的性能和速度。在200景约3 000像素×3 000像素大小，分辨率约8 m×8 m Google Earth测试图集上（其中，170景有机场，30景无机场），笔者提出的线段区域候选结合迁移学习深度CNN的识别方法可以得到相比其他传统方法明显的准确率提升，见表1，并且在GTX 780 GPU平台上，可以将耗时控制在秒级。

### 3.2.2 基于边缘尺度的飞行器目标候选区域提取

尽管机场识别为后续的飞行器识别给

出了有限的识别范围，但在Google Earth 19级图像上，典型的机场区域边长也可达上万像素。采用传统候选区域提取方法，也会得到百万数量级的候选区。为了提高识别效率，应尽可能减少待识别候选区域，并且尽量减少回归率的降低，结合飞行器目标尺度和边缘分布规律，笔者提出了一些候选区域优化方法：根据Google Earth 19级分辨率图像飞行器目标尺度分布范围，排除尺度过小或过大的候选区域；根据飞行器目标候选区域边缘密度分布特点，排除明显分布不符合目标特点的候选区域。该方法主要受Edgebox算法启发，能准确覆盖目标的候选区域，其边框与物体边缘的交叉机会较小，且区域内所含不同类型的目标，其边缘分布存在一定规律。如图6所示，在边缘图像中，不含目标的平滑区域其边缘密度相对较低，而复杂建筑区或者森林区域等边缘密度较高。因此，可以利用此先验来对飞行器目标候选区域进行筛选。

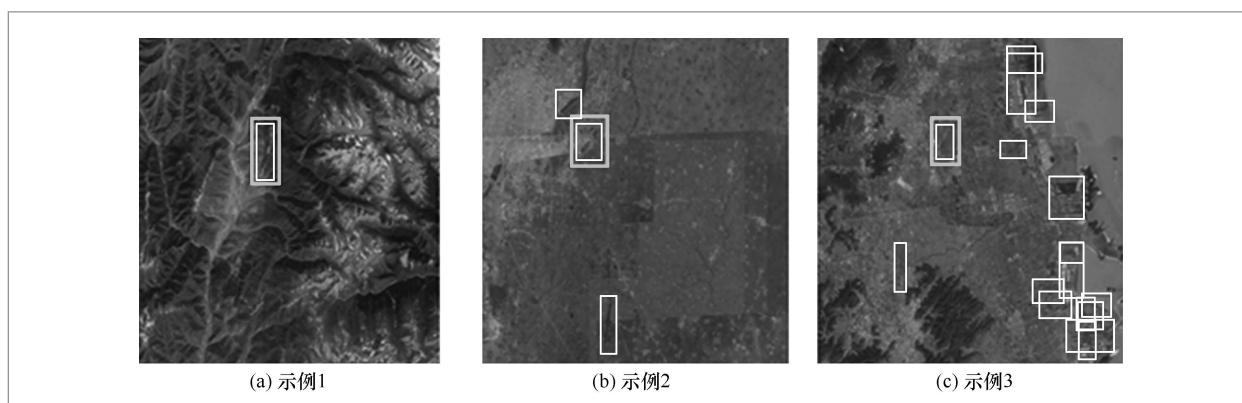


图5 基于边缘的机场候选区域提取结果示例

表1 基于所提线段候选方法的机场识别效果

候选方法	准确率	虚警率	平均耗时/s
基于显著性区域候选方法 <sup>[28]</sup> +CNN	68.5%	5%	3.85
基于显著性区域候选方法 <sup>[28]</sup> +SIFT	52%	8.5%	8.57
笔者提出的区域候选方法+ 8 纹理特征SVM <sup>[29]</sup>	62%	7.5%	68.64
笔者提出的区域候选方法+ SIFT	63%	15%	64.09
笔者提出的区域候选方法+ CNN	87%	12%	21.3

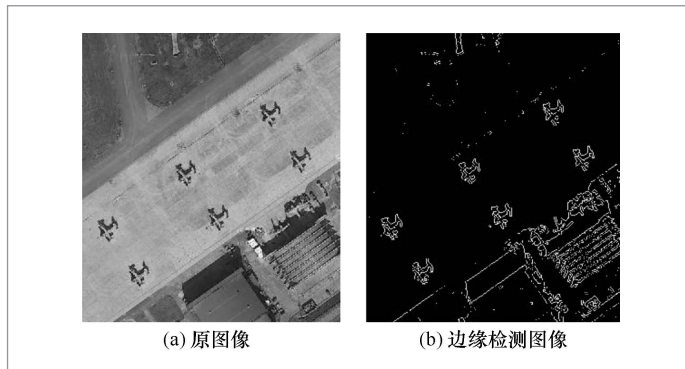


图6 利用 Prewitt 算子计算得到的边缘图像信息

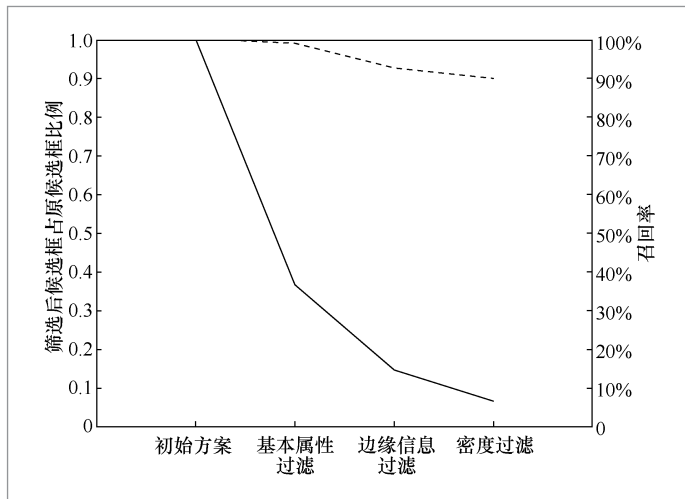


图7 过滤算法的效果及对应影响曲线

在实验中,该方法可以过滤平均90%的候选区域,只造成约10%的召回率下降,如图7所示。而根据应用侧重点的不同,可

以通过调整筛选阈值,选择相对保守或激进的筛选策略来调整实时性与准确性之间的平衡。

### 3.3 面向快速识别的级联式网络识别方法

即使采取了高效的候选区域筛选策略,对于“大范围、小目标”的遥感目标识别任务,若采用复杂的深度学习模型,其计算量仍然巨大。如前所述,采用候选筛选技术,可以将飞行器的候选区域减少90%,但对于机场图像中普遍初始有百万数量级的候选区域集合,其筛选后的区域数量仍有十万量级。即便采用GPU硬件加速手段,使用几十层的深度CNN也需几十分钟的识别时间。

为了解决采用深度模型进行精确预测与追求实时性的矛盾,基于遥感图像目标识别任务的特点,笔者设计了一个快速的、由粗到细的级联式CNN识别构架。该架构基于对生物视觉系统的模拟,不同输入对应不同的反映时间。简单的输入对应较短的判断时间,复杂的输入对应较长的处理时间。整体的识别框架如图8所示,级联式网络的基本思想采用准确率相对低但召回率相对较高的简单网络对候选区域进

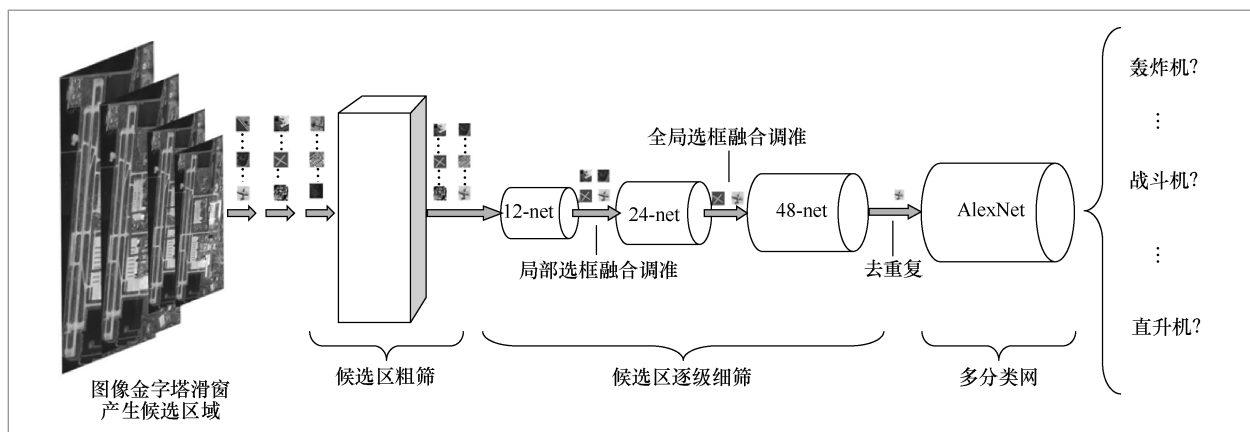


图8 面向快速飞行器识别的级联式 CNN 架构

行层层过滤。由于简单网络的计算代价较低,可以快速过滤一些容易判断的目标,而将少量难以判断的目标留给之后的复杂网络进行判断,从而在保持精度的前提下提高计算效率。该构架也是基于候选区域结合CNN识别的R-CNN架构,但与传统R-CNN采用单一的深层CNN方法不同,笔者精心设计了一系列CNN,并且以前后级联的方式连接。

在候选区域提取方法上,得益于之后级联网络的效率,可以选择覆盖全面的滑窗方法,并结合图像金字塔以提取不同尺度的目标。对于输入待识别图像,以1.414为缩放因子建立图像金字塔,然后以12像素 $\times$ 12像素大小的窗口,4个像素为步长对候选区域进行滑窗提取。由于固定12像素 $\times$ 12像素大小的窗口在不同的金字塔层对应着不同大小的区域,即可以用固定网络输入大小提取不同尺度的目标,避免了在固定图像上采用不同尺度的窗口提取区域时,为适应固定的CNN输入而造成的大量样本缩放开销,即用缩放整张图像来代替缩放单个

目标区域。

尺度滑窗后,再经过前述的候选区粗筛选,剩余约10万量级候选区域进入级联式CNN进行识别。如图9所示,笔者设计了4个CNN,由简单到复杂分别称之为12-net、24-net、48-net以及经过迁移学习的AlexNet。图9分别展示了12-net、24-net、48-net的网络结构。这3个结构逐渐复杂的网络被设计为判断是否为飞行器的二分类网络,其特征提取能力及识别能力也随网络层数的增加逐渐增强。输入候选区域经过级联网络逐层识别,在被某一网络判别为非飞行器目标即被丢弃,不再送入后续的深层次网络。在实验中,采用12-net即可在秒级过滤超过75%的候选区域。经过48-net处理后,候选区域剩余约1%剩余区域。最后得到有限的千余候选区域供深层次网络AlexNet进行进一步细分。由于候选区域大大减少,即使采用深度的AlexNet网络也可以实现秒级定位识别。

图10展示了级联式CNN在识别速度上

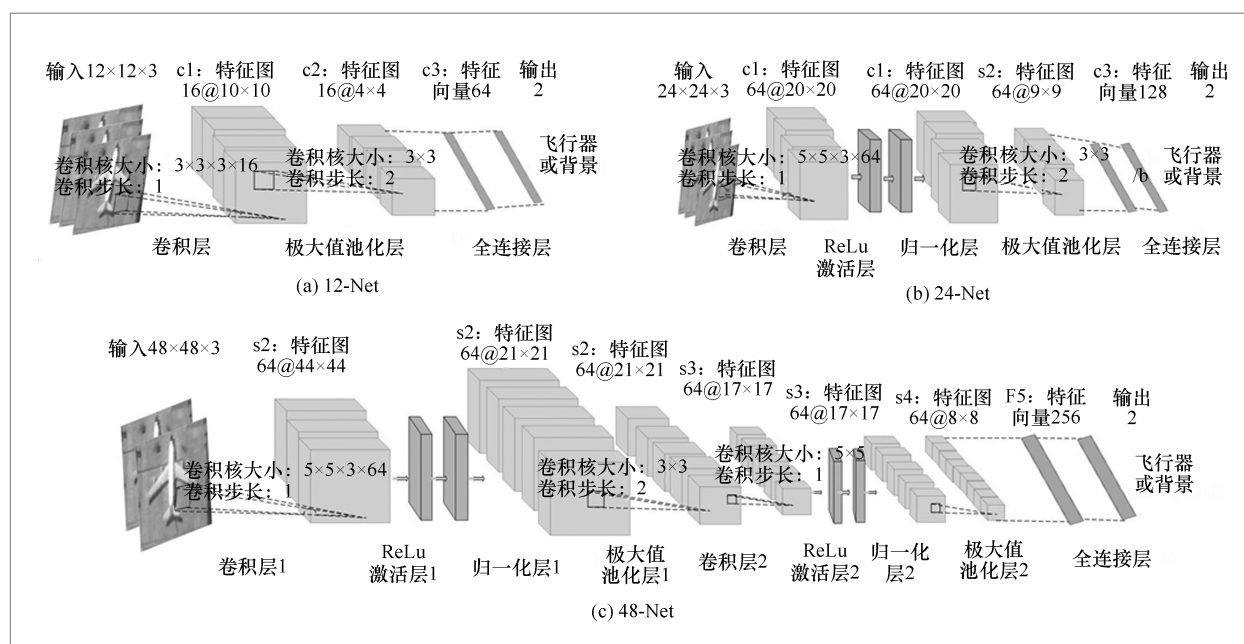


图9 12-net、24-net、48-net的网络结构

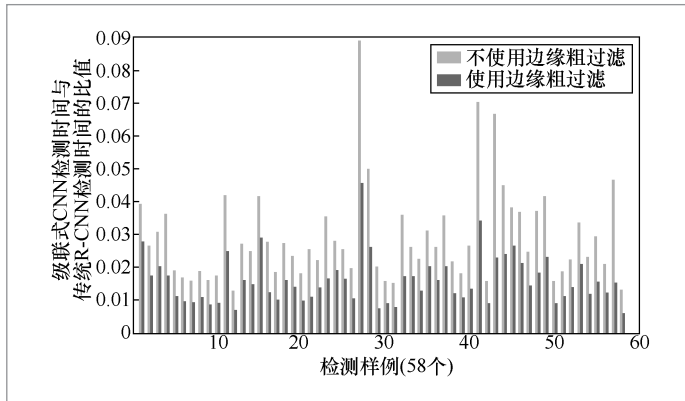


图 10 级联式 CNN 快速目标识别系统相对于传统 R-CNN 在识别时间上的对比

相对于传统R-CNN方法上的对比。图10中以R-CNN的识别时间为单位做了归一化。其中，横坐标代表试验中用到的58张机场图像。平均来看，笔者提出的级联式CNN

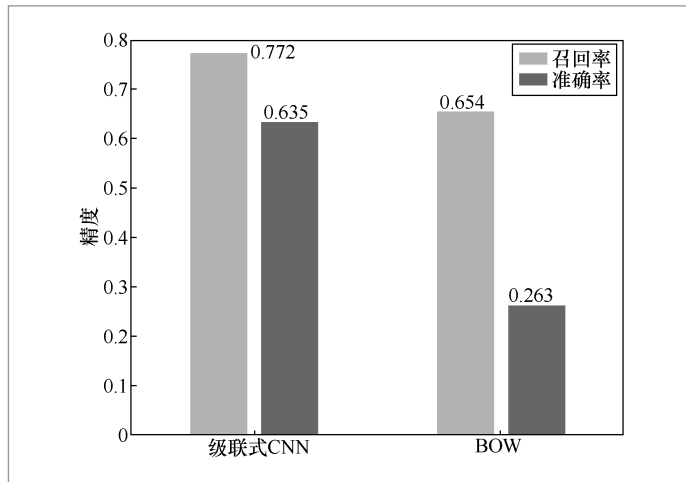


图 11 级联式 CNN 与 BOW R-CNN 在识别精度上的比较



图 12 级联式 CNN 飞行器目标识别结果展示 1

结构相对于传统R-CNN在速度上能提高40倍，结合基于图像低层边缘信息的“粗过滤”候选区域筛选方法，笔者提出的系统在速度上甚至可以提高100倍以上。而应用GTX 980 GPU，对于Google 19级机场区域，识别时间可以控制在几十秒。图11给出了采用级联CNN与基于BOW的传统方法在飞行器目标识别识别精度上的比较，表明了级联式CNN结构在高精度实时遥感目标识别应用中的可行性。图12、图13展示了部分飞行器识别结果。

## 4 结束语

深度学习技术为分析处理遥感大数据开辟了新的方向。应用深度模型可有效学习高分遥感图像中的复杂规律特征，从而提高检测识别等遥感应用的精度水平。在遥感样本采集相对困难的情况下，采用迁移学习等技术，可以快速构建高效的深度学习模型。结合应用目标的先验知识，可以提高深度学习的应用水平。针对实时性应用，采用级联式的深度学习网络可以进一步提升数据分析的精度和效率。

虽然深度学习技术在诸多遥感应用中取得了初步成效，但是也应该清醒地看到在其未来研究中仍然需要解决以下难题。

- 有标签样本相对稀缺与构建尽可能深网络的矛盾。遥感数据虽然从总量上看是大数据，但是对应到具体数据、具体应用，可采集的样本量就相对较少。因此，考虑无监督的训练方法以及可跨越数据类型的迁移学习方法可能会为此问题带来解决方案。

- 多源数据的融合问题。多源融合已成为目前遥感大数据的发展趋势之一。然而不同的数据类型其成像原理有很大差异，数据质量、结构差异明显，如何构建高效的深度学习模型，剔除各源数据的冗余信息，最

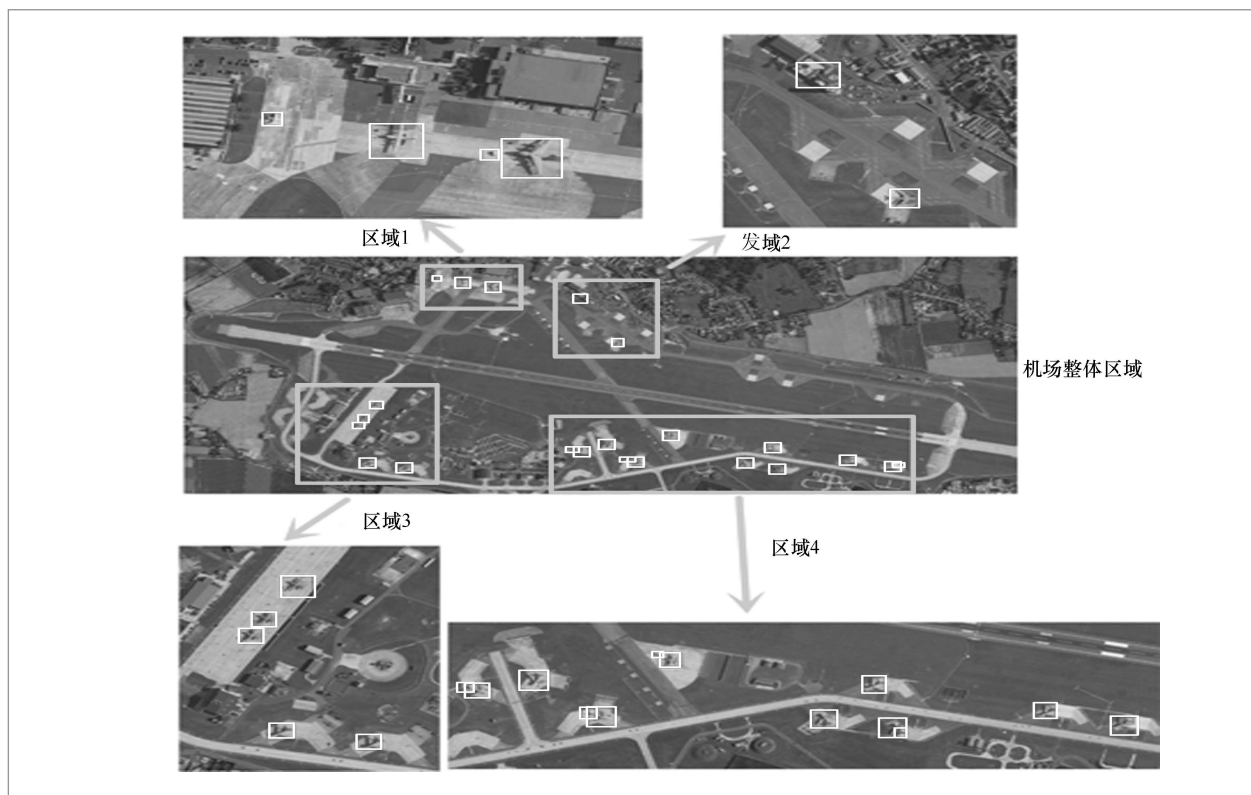


图13 级联式CNN飞行器目标识别结果展示2

大化信息互补也是深度模型在多源应用中的研究难点。此外,利用在异源数据上训练好的深度模型快速构建本源应用,将会极大减少多源应用的开发成本。

- 深度网络在遥感数据上的强化学习能力。虽然深度学习有提高应用效果的潜力,但是受限于遥感应用中可供训练的样本量以及不断变化的应用需求,需要研究具有不断强化能力的深度学习模型。在具体应用中可以只通过学习新添加的样本而不断提高模型的预测能力,并且能够对新引入类型目标进行强化学习,在较少影响以往预测能力的基础上,具备解决新问题的能力。

## 参考文献:

[1] 李国庆, 郭延辉. 大数据时代的对地观测科学研究[J]. 计算机学会通讯, 2013, 9(9): 27-31.

- LI G Q, WU Y H. Research on earth observation in age of big data[J]. Communications of CCF, 2013, 9(9): 27-31.
- [2] SUN H, SUN X, WANG H, et al. Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model[J]. IEEE Geoscience & Remote Sensing Letters, 2012, 9(1): 109-113.
- [3] ZHANG W, SUN X, FU K, et al. Object detection in high-resolution remote sensing images using rotation invariant parts based model[J]. IEEE Geoscience & Remote Sensing Letters, 2014, 11(1): 74-78.
- [4] BAI X, ZHANG H, ZHOU J. VHR object detection based on structural feature extraction and query expansion[J]. IEEE Transactions on Geoscience & Remote Sensing, 2014, 52(10): 6508-6520.
- [5] CHEN X, XIANG S, LIU C L, et al. Aircraft detection by deep belief nets[C]// Iapra Asian Conference on Pattern Recognition, November 5-8, 2013, Okinawa, Japan. New

- Jersey: IEEE Press, 2013: 54–58.
- [6] HAN J, ZHANG D, CHENG G, et al. Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2015, 53(6): 3325–3337.
- [7] HU X, SHEN J, SHAN J, et al. Local edge distributions for detection of salient structure textures and objects[J]. *IEEE Geoscience & Remote Sensing Letters*, 2013, 10(3): 466–470.
- [8] DIAO W, SUN X, ZHENG X, et al. Efficient saliency-based object detection in remote sensing images using deep belief networks[J]. *IEEE Geoscience & Remote Sensing Letters*, 2016, 13(2): 137–141.
- [9] ZHOU P, CHENG G, LIU Z, et al. Weakly supervised target detection in remote sensing images based on transferred deep features and negative bootstrapping[J]. *Multidimensional Systems & Signal Processing*, 2016, 27(4): 925–944.
- [10] SUN X, WANG H, FU K. Automatic detection of geospatial objects using taxonomic semantics[J]. *IEEE Geoscience & Remote Sensing Letters*, 2010, 7(1): 23–27.
- [11] INGLADA J, MICHEL J. Qualitative spatial reasoning for high-resolution remote sensing image analysis[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2009, 47(2): 599–612.
- [12] SUN H, SUN X, WANG H, et al. Automatic target detection in high-resolution remote sensing images using a contour-based spatial model[J]. *IEEE Geoscience & Remote Sensing Letters*, 2012, 9(5): 886–890.
- [13] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. *Neural Computation*, 1989, 1(4): 541–551.
- [14] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 23–28, 2014, Columbus, OH, USA. New Jersey: IEEE Press, 2014: 580–587.
- [15] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2014, 37(9): 1904–1916.
- [16] GIRSHICK R. Fast R-CNN[C]//The IEEE International Conference on Computer Vision, December 7–13, 2015, Chile. New Jersey: IEEE Press, 2015: 1440–1448.
- [17] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *Computer Science*, 2015(99): 1.
- [18] UIJLINGS J R R, SANDE K, GEVERS T, et al. Selective search for object recognition[J]. *International Journal of Computer Vision*, 2013, 104(2): 154–171.
- [19] ZITNICK C L, DOLL R P. Edge boxes: locating object proposals from edges[M]//Computer Vision–ECCV 2014. Switzerland: Springer International Publishing, 2014: 391–405.
- [20] CHENG M M, ZHANG Z, LIN W Y, et al. BING: binarized normed gradients for objectness estimation at 300fps[C]//IEEE Conference on Computer Vision and Pattern Recognition, June 23–28, 2014, Columbus, OH, USA. New Jersey: IEEE Press, 2014: 3286–3293.
- [21] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[J]. *Computer Science*, 2015: arXiv:1506.02640.
- [22] LIU W, ANGUOLOV D, ERHAN D, et al. SSD: single shot multibox detector[J]. *Computer Science*, 2015: arXiv:1512.02325.
- [23] HINTON G E. Deep belief networks[J]. *Scholarpedia*, 2009, 4(6): 804–786.
- [24] YU Y, GUAN H, JI Z. Rotation-invariant object detection in high-resolution satellite imagery using superpixel-based deep hough forests[J]. *IEEE Geoscience & Remote Sensing Letters*, 2015, 12(11): 1.
- [25] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. *Advances in Neural Information Processing Systems*, 2012, 25(2): 1–9.

- [26] MCCALLUM. BOW: a toolkit for statistical language modeling, text retrieval, classification and clustering[J]. Machine Learning, 1996 (1): 405-408.
- [27] OGAWA S, HIRAO K, HONGU T, et al. Line segment detector: US, US 20020154121 A1[P]. 2002.
- [28] ZHU D, WANG B, ZHANG L. Airport target detection in remote sensing images: a new method based on two-way saliency[J]. IEEE Geoscience & Remote Sensing Letters, 2015, 12(5): 1096-1100.
- [29] QU Y, LI C, ZHENG N. Airport detection base on support vector machine from a single image[C]// Fifth International Conference on Information, Communications and Signal Processing, December 6-9, 2005, Bangkok, Thailand. New Jersey: IEEE Press, 2005: 546-549.

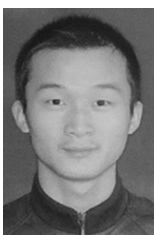
## 作者简介



牛新(1983-),男,博士,国防科学技术大学并行与分布处理重点实验室助理研究员,主要研究方向为计算机系统结构、遥感图像处理、机器学习。



钱勇(1966-),男,博士,国防科学技术大学并行与分布处理重点实验室常务副主任、研究员、博士生导师,计算机系统结构学科方向学术带头人,担任中国计算机学会体系结构专家委员会主任,是国家自然科学基金杰出青年基金获得者,入选教育部新世纪优秀人才支持计划。主要研究方向为计算机系统结构、并行计算、计算机应用。



张鹏(1993-),男,国防科学技术大学并行与分布处理重点实验室硕士生,主要研究方向为计算机系统结构、遥感图像处理、机器学习。



曹玉社(1990-),男,国防科学技术大学并行与分布处理重点实验室硕士生,主要研究方向为计算机系统结构、遥感图像处理、机器学习。

收稿日期: 2016-08-20

基金项目: 国家自然科学基金资助项目(No.U1435219, No.61402507)

Foundation Items: The National Natural Science Foundation of China (No.U1435219, No.61402507)